ANNEX A · PRACTICAL 4.7 — SMALLER, SECRET, UNCHANGED

NAME _____ DATE _____

Tramuntana Apps must send its weekly archive to the accountant: it should be **small**, nobody else may **read** it, and she must know it arrived **unchanged**. Three jobs, three kinds of tool. You install them, measure them and choose. Work in `deb-lab`.

TASKS 1–4 · 20 MIN

TASK 1 Install the tools, build the sample

```
sudo apt install zstd zip unzip age
mkdir -p ~/lab37/sample
cd ~/lab37
seq 1 4000000 > sample/numbers.txt
cp -r /usr/share/common-licenses sample/licences
head -c 20M /dev/urandom > sample/random.bin
tar -cf sample.tar sample
ls -lh
```

`seq` writes the numbers from 1 to 4,000,000: 30 MB of text. `tar -cf` packs the folder into one file, **without** compressing. Size of `sample.tar`: _____

DONE WHEN `sample.tar` is about 50M.

TASK 2 The race

Same file, four compressors. `time` measures the command (read the line `real`); `-k` keeps the original.

```
time gzip -k sample.tar
time zstd -k sample.tar
time xz -k sample.tar
time zip -q -r sample.zip sample
ls -l
```

Tool	File it made	Size (MB)	Time `real` (s)
`gzip`			
`zstd`			
`xz`			
`zip`			

`ls -l` shows bytes: 29589180 is 29.6 MB. Smallest: _____ Fastest: _____

DONE WHEN four rows with a size and a time.

TASK 3 Why not smaller?

Compress the two big files of the sample on their own:

```
gzip -k sample/numbers.txt sample/random.bin
ls -l sample
rm sample/*.gz
```

	Before (bytes)	After `gzip` (bytes)
`numbers.txt`		
`random.bin`		

One of them **grew**. Why can it not be compressed?

_____

_____

DONE WHEN you used the word “repetition” (or “pattern”) in your answer.

TASK 4 Nothing was lost

Unpack the `.zst` archive into another folder and compare with the original:

```
mkdir out
tar -xf sample.tar.zst -C out
diff -r sample out/sample
```

What does `diff` print? _____

DONE WHEN `diff` is silent: compression gave back every byte.

TASKS 5–7 · 14 MIN

TASK 5 A secret

The Wi-Fi password of the office should not travel as plain text. Encrypt it with a passphrase (`-p`). Choose one and **type it**: do not leave it empty.

```
cp ~/tramuntana-share/admin/wifi.txt .
cat wifi.txt
age -p -o wifi.txt.age wifi.txt
file wifi.txt.age
```

What does `file` say it is? _____

Now decrypt (`-d`) and compare. Then try once with a wrong passphrase.

```
age -d -o wifi-back.txt wifi.txt.age
diff wifi.txt wifi-back.txt
age -d wifi.txt.age
```

Error with the wrong passphrase:

_____

DONE WHEN `wifi-back.txt` is identical to `wifi.txt`, and the wrong passphrase opened nothing.

TASK 6 Unchanged?

Write the fingerprints of the two files you would send, then check them (`-c`):

```
sha256sum sample.tar.zst wifi.txt.age > SHA256SUMS
cat SHA256SUMS
sha256sum -c SHA256SUMS
```

How many characters has one fingerprint? _____
Result of the check: _____
Now somebody adds **one** byte to a file on the way (`>>` adds to the end):

```
echo x >> wifi.txt.age
sha256sum -c SHA256SUMS
```

Copy the line of the damaged file:

_____

DONE WHEN one file says `OK` and the other says `FAILED`.

TASK 7 Choose

Use **your** table of task 2.

Every night Tramuntana compresses 100 GB of projects. Which tool, and why?

_____

_____

The archive goes to a client who uses Windows and installs nothing. Which tool?

_____

You must compress **and** encrypt. In which order? (Think of task 3: what does an encrypted file look like?)

_____

_____

```
cd ~
rm -r ~/lab37
```

DONE WHEN three answers with a reason, and `~/lab37` is deleted.

Hand in on Classroom, by Monday 08/02 at 23:59: `practical-4.7-surname.md`. Your tables of tasks 2 and 3, the evidence of tasks 5 and 6 (the error and the `FAILED` line), and your three answers of task 7.

Marking (10): tools installed and table of task 2 complete 3 · task 3 explained 2 · tasks 5–6 with evidence 2 · task 7: three choices with reasons from your own numbers 2 · clear Markdown 1.

EXIT TICKET

1. Why does a file of random bytes not get smaller with `gzip`?

- ☐ **A** It is too big
- ☐ **B** It is already encrypted
- ☐ **C** It has no repetition to describe in a shorter way
- ☐ **D** `gzip` only works with text

2. Which tool makes a file unreadable without a passphrase?

- ☐ **A** `age`
- ☐ **B** `zstd`
- ☐ **C** `sha256sum`
- ☐ **D** `tar`

3. `sha256sum -c SHA256SUMS` answers `report.pdf: FAILED`. What do you know?

- ☐ **A** The file is encrypted
- ☐ **B** The file is too big
- ☐ **C** The file was compressed twice
- ☐ **D** The file is not the same as when the list was written

4. A client with Windows, who installs nothing, needs a folder of text logs in one small file. You send…

- ☐ **A** `logs.tar.zst`
- ☐ **B** `logs.zip`
- ☐ **C** `logs.tar`
- ☐ **D** `logs.age`